X-ray crystallographic studies on the hydrogen isotope effects of green fluorescent protein at sub-ångström resolutions

AUTHOR(S):

Tai, Yang; Takaba, Kiyofumi; Hanazono, Yuya; Dao, Hoang-Anh; Miki, Kunio; Takeda, Kazuki

CITATION:

Tai, Yang ...[et al]. X-ray crystallographic studies on the hydrogen isotope effects of green fluorescent protein at sub-ångström resolutions. Acta Crystallographica Section D Structural Biology 2019, 75(12): 1096-1106

ISSUE DATE:

2019-12

URL:

http://hdl.handle.net/2433/245188

Acta Cryst D STRUCTURAL BIOLOGY
ISSN 2059-7983

X-ray crystallographic studies on the hydrogen isotope effects of green fluorescent protein at sub-ångström resolutions

Yang Tai, Kiyofumi Takaba,‡ Yuya Hanazono,§ Hoang-Anh Dao, Kunio Miki and Kazuki Takeda*

Department of Chemistry, Graduate School of Science, Kyoto University, Sakyo-ku, Kyoto 606-8502, Japan. *Correspondence e-mail: ktakeda@kuchem.kyoto-u.ac.jp

Received 31 July 2019
Accepted 28 October 2019

Edited by R. McKenna, University of Florida, USA

‡ Present address: Biostructural Mechanism Laboratory, RIKEN SPring-8 Center, 1-1-1 Kouto, Sayo, Hyogo 679-5148, Japan.
§ Present address: Graduate School of Information Sciences, Tohoku University, Aoba-ku, Sendai, Miyagi 980-8579, Japan.

Keywords: green fluorescent protein; GFP; hydrogen bonding; perdeuterated protein; pH/pD dependence; protonation.

PDB references: S65T/F99S/M153T/V163A variant of green fluorescent protein, perdeuterated, pD 8.5, 6kkz; perdeuterated, pD 7.0, 6kl0; nondeuterated, pD 8.5, 6kl1

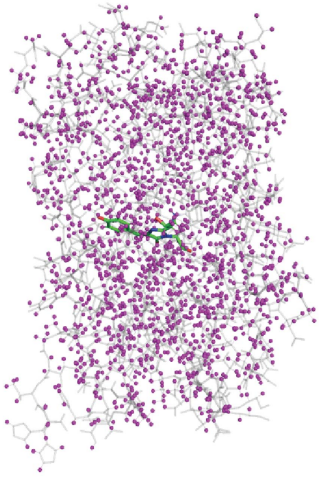
Supporting information: this article has supporting information at journals.iucr.org/d

Hydrogen atoms are critical to the nature and properties of proteins, and thus deuteration has the potential to influence protein function. In fact, it has been reported that some deuterated proteins show different physical and chemical properties to their protiated counterparts. Consequently, it is important to investigate protonation states around the active site when using deuterated proteins. Here, hydrogen isotope effects on the S65T/F99S/M153T/V163A variant of green fluorescent protein (GFP), in which the deprotonated B form is dominant at pH 8.5, were investigated. The pH/pD dependence of the absorption and fluorescence spectra indicates that the protonation state of the chromophore is the same in protiated GFP in H_2O and protiated GFP in D_2O at pH/pD 8.5, while the pK_a of the chromophore became higher in D_2O. Indeed, X-ray crystallographic analyses at sub-ångström resolution revealed no apparent changes in the protonation state of the chromophore between the two samples. However, detailed comparisons of the hydrogen OMIT maps revealed that the protonation state of His148 in the vicinity of the chromophore differed between the two samples. This indicates that protonation states around the active site should be carefully adjusted to be the same as those of the protiated protein when neutron crystallographic analyses of proteins are performed.

1. Introduction

Hydrogen atoms play important roles in the mechanisms of protein function, especially for enzymes and photoactive proteins (Schowen *et al.*, 2000; Ishikita & Saito, 2013). Structural information on H atoms around the active site is thus crucial to understanding the mechanisms behind the activities of the protein. Isotope effects on physical and chemical properties have been investigated for various proteins. Consequently, the properties and reactivities of many deuterated proteins have been shown to differ from those of their protiated counterparts (Hattori *et al.*, 1965; Chen *et al.*, 1984; Brockwell *et al.*, 2001; Cioni & Stambini, 2002; Cleland, 2005; Schramm, 2007; Piszczek *et al.*, 2011; Gu & Zhang, 2013).

X-ray crystallography is one of the most widely used methods for the structural analysis of proteins. Although this method is very effective for determining the atomic coordinates of non-H atoms, the determination of the precise coordinates of H atoms using this approach is very laborious. Therefore, neutron crystallography is often used as an alternative approach for the determination of the coordinates of H atoms (reviewed in Chen & Unkefer, 2017; Oksanen *et al.*, 2017; Ashkar *et al.*, 2018). Furthermore, the dissociation and



polarization of H atoms can be addressed in detail by using both X-ray and neutron data at high resolutions, because the distribution of the electron density is different from that of the nucleus. However, absorption and inelastic scattering by protium atoms decrease the diffraction data quality in neutron crystallography. Deuteration of proteins by soaking in D_2O or the use of completely deuterated (perdeuterated) proteins are standard procedures in neutron diffraction measurements in order to collect diffraction data at the highest possible resolution. Because the soaking method can replace only a portion of the exchangeable H atoms, large crystals of at least 1 mm^3 are used to collect neutron data to ∼2.0 Å resolution for the precise analysis of H atoms (Afonine *et al.*, 2010; Blakeley *et al.*, 2015). By using perdeuterated proteins, the smallest size of crystals previously reported for neutron data collection to better than ∼2.0 Å resolution was ∼0.2 mm^3 (Hazemann *et al.*, 2005; Chen & Unkefer, 2017). However, it is crucial to study the effects of deuteration on protein structures, especially on the active sites.

Green fluorescent protein (GFP), which was discovered from the jellyfish *Aequorea victoria* (Shimomura *et al.*, 1962), consists of 238 amino acids. The structure of GFP consists of an 11-stranded β-barrel plugged by a chromophore (Yang *et al.*, 1996; Ormö *et al.*, 1996; Fig. 1*a*). The chromophore is synthesized from three intrinsic residues, Ser65, Tyr66 and Gly67, by post-translational reactions. Two forms of the chromophore, the 'A' and 'B' forms, with different protonated states exist in wild-type GFP (wt-GFP) in a population ratio of 4:1–6:1 (Brejc *et al.*, 1997; Chattoraj *et al.*, 1996). In the A form the phenolic group of Tyr66 in the chromophore is neutral. On the other hand, in the B form this group is deprotonated and anionic (Heim & Tsien, 1996; Chattoraj *et al.*, 1996; Kummer *et al.*, 1998; Fig. 1*b*). The A and B forms can be individually stabilized by mutations of residues around the chromophore. The population ratio of the B form is increased in some variants such as S65T and E222Q, in which the hydrogen-bond relay between O^{η} and N^2 is disrupted (Brejc *et al.*, 1997; Elsliger *et al.*, 1999; Fig. 1*b*). Such variant GFPs have higher emission intensities than wt-GFP (Heim *et al.*, 1995; Heim & Tsien, 1996). Therefore, these variants are useful for applications in various fields, for example cytology, medical science and biological science (Sirerol-Piquer *et al.*, 2012; Morris, 2013). However, the variants have a large pH dependence of their visible absorption and fluorescence (Kneen *et al.*, 1998; Elsliger *et al.*, 1999). Moreover, wt-GFP has been shown to have a prolonged fluorescence lifetime after H/D exchange (McAnaney *et al.*, 2002). These findings demonstrate that the fluorescence mechanism is closely related to the H atoms around the chromophore.

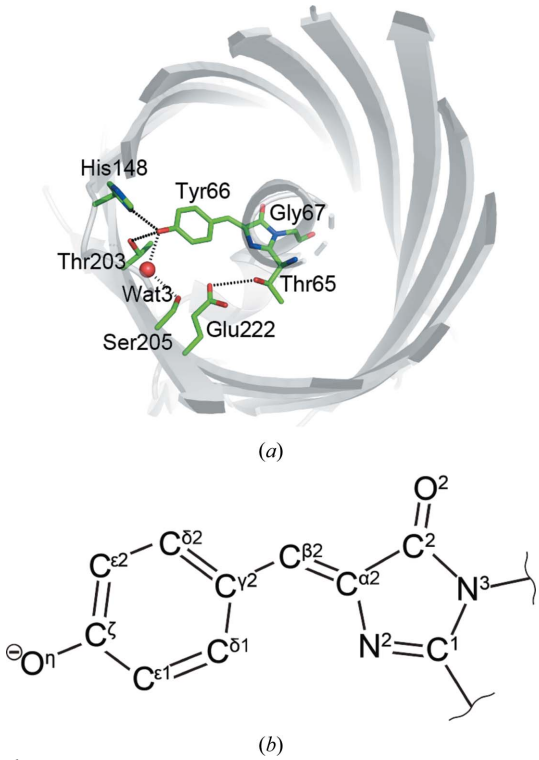


Figure 1
Structure of GFP from *A. victoria*. (*a*) A ribbon model of h-GFP_{C3} (PDB entry 6jgi) is viewed from the top of the β-barrel. (*b*) The structural formula of the anionic B form of the chromophore. The O^{η} atom in the tyrosyl group is deprotonated and negatively charged in the B form.

In our previous high-resolution X-ray analyses of the A and B forms using variants of GFP, the H atoms of most residues and some internal waters were observed clearly, and the results enabled us to discuss the differences in the hydrogen-bond networks of the A and B forms around the chromophore (Takaba *et al.*, 2019). According to the pH-dependence of tthe UV–Vis spectra, almost all molecules adopt the major form for the respective variants at pH 8.5. Additionally, the spectroscopic properties at pH 8.5 are the same as those at the physiological pH of ∼7. Therefore, GFP crystals were kept at pH 8.5 during X-ray analyses for all variants. However, there is no guarantee that deuterated GFP will show an identical protonation state at the same pH/pD. Neutron analyses will be needed to obtain more detailed pictures of the H atoms in GFP, but such analyses must include an estimation of the influence of deuteration. To date, no atomic-level structural details of the influence of deuteration have been reported for GFP, although neutron scattering analyses of deuterated GFP have previously been published (Nickels *et al.*, 2012, 2013). In the present study, therefore, we compared deuterated and protiated samples of a GFP variant by means of spectroscopic and high-resolution X-ray crystallographic analyses.

2. Materials and methods

2.1. Protein expression and purification

Protiated samples of the S65T/F99S/M153T/V163A variant (h-GFP_{C3}) containing the cycle3 (F99S/M153T/V163A) mutation (Fukuda *et al.*, 2000) were prepared as described previously (Takaba *et al.*, 2019). Deuterated samples of this variant (d-GFP_{C3}) were prepared as follows. Firstly, transformed *Escherichia coli* BL21(DE3) pLysS cells harbouring a pET-21a-based plasmid were grown overnight in 4 ml LB medium containing 100 µg ml^{-1} ampicillin and 33 µg ml^{-1} chloramphenicol at 37°C. After centrifugation, the harvested

cells were transferred into 400 ml deuterated medium consisting of 40 ml Bioexpress Cell Growth Media (U-D, 98%) (Cambridge Isotope Laboratories) and 360 ml D_2O (99.8%) (Euriso-Top). The cells were grown at 37°C for about 6 h. At an OD_{600} of ∼0.7, expression was induced by the addition of 3 m*M* IPTG dissolved in D_2O. The cells were cultured for a further 24 h at 22°C. After harvesting, subsequent purifications were performed using protiated buffer solutions in the same way as for the protiated samples (Takaba *et al.*, 2019). Briefly, the collected cells were suspended in lysate buffer consisting of 200 m*M* Tris–HCl pH 8.5 and BugBuster (Novagen) and were shaken for 24 h at room temperature. The solution separated by centrifugation was purified using an Ni–NTA affinity column (Qiagen). After removal of the His tag using subtilisin, the sample was purified using an anion-exchange Mono Q column (GE Healthcare). The purified samples of d-GFP_{C3} were incubated at 4°C for more than 3 d before their use in further experiments. Protiated (h-GFP_{C0}) and deuterated (d-GFP_{C0}) samples of the S65T variant without the cycle3 mutation were prepared in the same way as the h-GFP_{C3} and d-GFP_{C3} samples, respectively.

2.2. Mass spectrometry

A matrix solution of D_2O containing 10 mg ml^{-1} protiated sinapinic acid, 0.1%(*v*/*v*) deuterated trifluoroacetic acid and 50%(*v*/*v*) acetonitrile in D_2O was prepared for the deuterated samples. The deuterated protein (0.1 mg ml^{-1} in D_2O) was mixed with the matrix solution in a mixing ratio of 1:1. 0.5 µl of the mixed solutions was then naturally dried for over 30 min on a sample plate at room temperature. The masses were measured by matrix-assisted laser desorption ionization time-of-flight mass spectrometry with a Voyager-DE RP (Applied Biosystems). Errors were calculated as the standard deviation of 9–13 independent measurements. The calculated mass of h-GFP_{C0} with the His tag (25 600 Da) was used as a calibration standard. The masses of protiated samples were measured in the same way using protiated matrix solution.

2.3. Measurement of the pD dependence of absorption spectra

The pD dependences of the UV–Vis absorption spectra of h-GFP_{C3} and d-GFP_{C3} were measured in D_2O solution containing 100 m*M* NaCl and 0.1 mg ml^{-1} protein. The pD values of the samples were set in the range 4.0–8.5 in intervals of 0.5 using a wide-range buffer series consisting of 10 m*M* MES, 10 m*M* MOPS, 10 m*M* citrate and NaOD (Kneen *et al.*, 1998). The absorbance from 250 to 550 nm was measured at 20°C using a V-630 spectrophotometer (JASCO). The pD values were checked with a B-212 pH meter (HORIBA) immediately after the spectroscopic measurements. The pD values were calibrated using the formula pD = pH + 0.4, where pH is the value measured with the pH meter (Glasoe & Long, 1960). The pH dependence of h-GFP_{C3} in H_2O was measured in the same way. The pK_a values were determined by fitting to the Henderson–Hasselbalch equation. The pK_a values of d-GFP_{C3} and h-GFP_{C3} in D_2O were calculated using the calibrated pD values. The presence of the intermediate 'I' form was ignored in the calculation, because the A form of the S65T variant and its related variants show extremely low fluorescence, implying inconvertibility to the I form (Elsliger *et al.*, 1999).

2.4. Crystallization

Crystals of d-GFP_{C3} (pD 8.5) were prepared in almost the same manner as those of h-GFP_{C3} (pH 8.5) (Takaba *et al.*, 2019). Briefly, microseed crystals were prepared from h-GFP_{C3} by the hanging-drop vapour-diffusion method at 35°C. The sample solution (∼10 mg ml^{-1} h-GFP_{C3}, 10 m*M* Tris–HCl pH 8.5 dissolved in H_2O) was mixed with a precipitant solution consisting of 20%(*w*/*v*) PEG 4000, 25 m*M* $MgCl_2$, 20 m*M* Tris–HCl pH 8.5 dissolved in H_2O at a 1:1 ratio. Clusters of needle-like microcrystals were obtained in ∼1 week. A suspended solution of the crushed crystals was serially diluted by adding the precipitant solution and was used as a microseed solution. Microseeding was performed by the sitting-drop vapour-diffusion method at 35°C. 0.3 µl seed solution was added to 90 µl crystallization solution consisting of 5 mg ml^{-1} d-GFP_{C3}, 5%(*w*/*v*) PEG 4000, 12.5 m*M* $MgCl_2$, 10 m*M* Tris–DCl pD 8.5 dissolved in D_2O. The solution was equilibrated against 500 µl reservoir solution [10%(*w*/*v*) PEG 4000, 25 m*M* $MgCl_2$, 20 m*M* Tris–DCl pD 8.5 dissolved in D_2O]. Single crystals of typically 0.2–0.3 mm in length were obtained ∼1 week after microseeding. Crystals with a low aspect ratio and a good appearance were used as macroseeds after washing with the reservoir solution. The seed crystals were grown at 35°C in 200 µl crystallization solution, which was equilibrated against 500 µl reservoir solution. The crystals reached 1.3 × 0.3 × 0.1 mm in size three weeks after macroseeding. Crystals of d-GFP_{C3} (pD 7.0) and h-GFP_{C3} (pD 8.5) were prepared in the same way to those of d-GFP_{C3} (pD 8.5).

2.5. X-ray data collection and reduction

A crystal of d-GFP_{C3} (pD 8.5) was successively soaked in a series of cryoprotectant solutions in which the amount of PEG 4000 was increased in a stepwise manner; the final solution consisted of 40%(*w*/*v*) PEG 4000, 25 m*M* $MgCl_2$, 20 m*M* Tris–DCl pD 8.5. The treated crystals were flash-cooled in a nitrogen-gas stream at 100 K. The other crystals were cryoprotected in a similar way. For the crystallographic analyses in this paper, the pD values for the stock solutions of buffers are referenced, because the shifts in the final cryoprotectant solutions were adjusted within 0.3 pD units.

Diffraction data for d-GFP_{C3} (pD 8.5) were collected on the BL41XU beamline at SPring-8, while those for d-GFP_{C3} (pD 7.0) and h-GFP_{C3} (pD 8.5) were collected on the BL44XU beamline at SPring-8. The wavelength of the incident X-rays for d-GFP_{C3} (pD 8.5) was set to 0.70 Å and the diffraction intensities were measured using a PILATUS 6M detector (Dectris). The wavelength of the incident X-rays was set to 0.75 Å and an MX300HE detector (Rayonix) was used for data collection from the d-GFP_{C3} (pD 7.0) and h-GFP_{C3} (pD 8.5) crystals. The crystal of d-GFP (pD 8.5) was cooled during

Table 1
Data-collection and crystallographic statistics.

Values in parentheses are for the outer shell.

	d-GFP$_{C3}$ (pD 8.5)	d-GFP$_{C3}$ (pD 7.0)	h-GFP$_{C3}$ (pD 8.5)
Diffraction source	BL41XU, SPring-8	BL44XU, SPring-8	BL44XU, SPring-8
Wavelength (Å)	0.70	0.75	0.75
Temperature (K)	100	50	50
Detector	Dectris PILATUS-6M	Rayonix MX300HE	Rayonix MX300HE
Crystal-to-detector distance (mm)	210	80/80/80†	80/—/80†
Rotation range per image (°)	0.5	0.5/0.5/0.5†	0.5/—/0.5†
Total rotation range (°)	360	180/180/180†	180/—/180†
Exposure time per image (s)	0.5	0.5/0.5/0.5†	0.5/—/0.5†
Dose per position (Gy)	2×10^4	3.6×10^4/3.9×10^4/3.7×10^4†	3.6×10^4/—/3.9×10^4†
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
a, *b*, *c* (Å)	50.79, 62.17, 69.03	50.83, 62.40, 68.88	50.91, 62.11, 68.86
Mosaicity (°)	0.16–0.24	0.14–0.43	0.14–0.40
Resolution range (Å)	50–0.90 (0.92–0.90)	50–0.80 (0.81–0.80)	50–0.85 (0.86–0.85)
Total No. of reflections	916251	2662411	1602184
No. of unique reflections	159331	229145	189497
Completeness (%)	98.3 (80.3)	100 (99.8)	99.3 (98.0)
Multiplicity	5.8 (3.0)	11.6 (6.0)	8.5 (6.2)
$CC_{1/2}$‡	(0.473)	(0.475)	(0.526)
$\langle I/\sigma(I)\rangle$	18.0 (1.0)	29.6 (1.2)	21.6 (1.3)
R_{merge}§ (%)	9.4 (83.9)	9.2 (145.9)	8.2 (125.8)
Wilson *B* factor (Å^2)	4.7	5.0	4.4

† Values are given for high-resolution/medium-resolution/low-resolution data. ‡ $CC_{1/2}$ values are calculated by splitting the data randomly into half data sets. § $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$.

data collection by a nitrogen-gas stream at 100 K, while the crystals of d-GFP (pD 7.0) and h-GFP (pD 8.5) were cooled during data collection by a helium-gas stream at 50 K. The helical data-collection method was applied (Flot *et al.*, 2010). The X-ray absorption dose was calculated with *RADDOSE* (Paithankar *et al.*, 2009). High-resolution data were measured using the helical data-collection method, while medium/low-resolution data were separately measured from non-irradiated portions of the crystal. The details of these measurement conditions are summarized in Table 1. Diffraction data sets were processed and scaled with the *HKL*-2000 program suite (Otwinowski & Minor, 1997). The resolution limits were defined at a $CC_{1/2}$ of ∼0.5 (Karplus & Diederichs, 2012). The crystallographic statistics are listed in Table 1.

2.6. Structure refinement

The initial stages of the refinement calculations were carried out using *Phenix* (Adams *et al.*, 2010). The geometric restraint for the chromophore was generated from the structure of h-GFP$_{C3}$ and was gradually reduced during the course of refinement (Takaba *et al.*, 2019). The structures were manually corrected using *Coot* (Emsley *et al.*, 2010) by monitoring the $2F_{obs} - F_{calc}$ and $F_{obs} - F_{calc}$ maps. All deuterium atoms were added to the models as riding hydrogens, and were treated as protium atoms during the refinement calculations. Some final steps of the refinement were performed using *SHELXL* (Sheldrick & Schneider, 1997). Only H atoms that were confirmed in the $F_{obs} - F_{calc}$ OMIT map at the 1.5σ contour level were included in the final models. Errors in the interatomic distances were estimated as the standard deviations given by full-matrix least-squares refinement with *SHELXL* after removing all of the restraints. Based on plots of bond-length estimated standard deviations *versus* average equivalent *B* values (B_{eq}) for our structures, it was confirmed that the values of the estimated standard deviations are not saturated and are not restrained for atoms with large B_{eq} (Supplementary Fig. S1). Differences between the same hydrogen-bond distances as well as the covalent bond lengths (l_1 and l_2) of two structures were evaluated by the value of the σ level, which is calculated using $|l_1 - l_2|/(\sigma_{l1}^2 + \sigma_{l2}^2)^{1/2}$. The refined structures were validated with *MolProbity* (Chen *et al.*, 2010). Figures showing molecular models were prepared using *PyMOL* (DeLano, 2002), while electron-density maps were calculated with *Phenix*. The anisotropy, which is the ratio of the smallest to the largest eigenvalue of the atomic displacement parameters, was calculated with *PARVATI* (Merritt, 1999). The hydrogen visualization ratio for each residue was calculated as N_{obs}/N_{sum}, where N_{obs} is the sum of the occupancies of modelled H atoms and N_{sum} is the theoretical sum of the occupancies of H atoms. The accessible surface area (ASA) ratio for each residue was calculated with the 'Accessible Surface Area and Accessibility Calculation for Protein' tool (Center for Informational Biology, Ochanomizu University, Japan; http://cib.cf.ocha.ac.jp/bitool/ASA/). The free ASA for various residues in solution used values from Lins *et al.* (2003). The ASA ratio of residue *i* was calculated as (ASA$_i$ in GFP)/(free ASA$_i$ in solution).

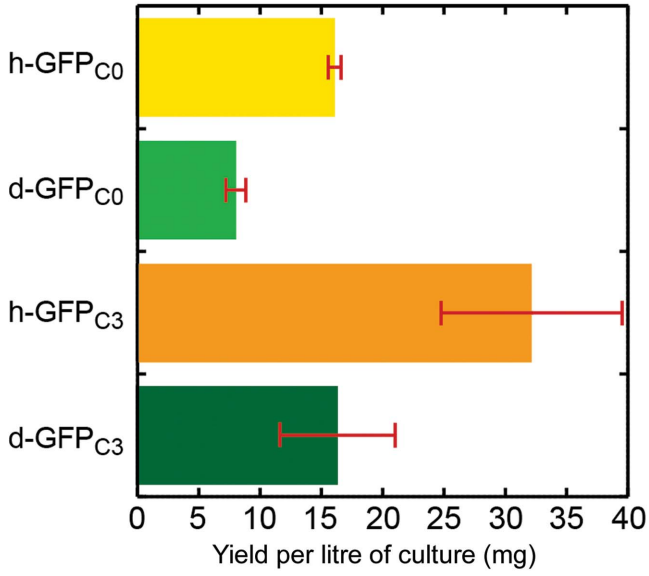
3. Results

3.1. Preparation of perdeuterated GFP

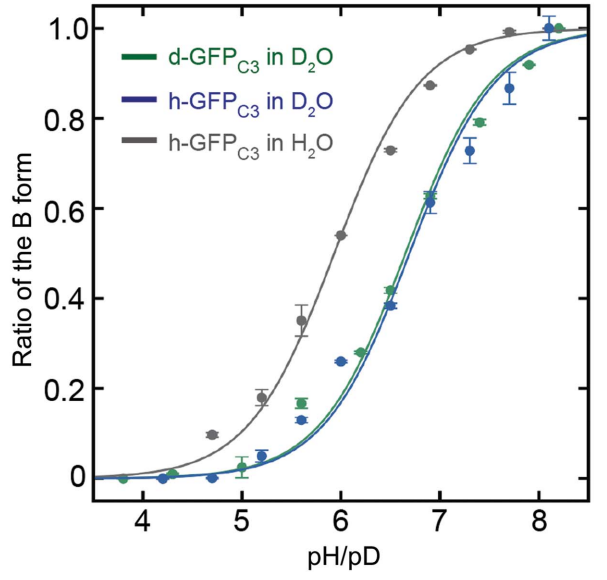
At the beginning of this study, we investigated the change in yield on deuteration. The yield of d-GFP$_{C0}$ was 8.0 ± 0.8 mg per litre of culture using perdeuterated medium, while that of h-GFP$_{C0}$ was 16.3 ± 0.5 mg per litre of culture (Fig. 2*a*). The yield of the deuterated sample was reduced to half of that of the protiated sample, as reported for wt-GFP (Hohlefelder *et al.*, 2013). On introducing the cycle3 mutation (Fukuda *et al.*, 2000), the yields of h-GFP$_{C3}$ and d-GFP$_{C3}$ were increased twofold to 16.1 ± 4.8 mg per litre of culture for d-GFP$_{C3}$ and 32.0 ± 7.5 mg per litre of culture for h-GFP$_{C3}$. The molecular masses of d-GFP$_{C0}$ and h-GFP$_{C0}$ were determined to be 27 752 ± 24 and 26 036 ± 4 Da, respectively. Those of d-GFP$_{C3}$ and h-GFP$_{C3}$ were determined to be 27 509 ± 19 and 25 857 ± 12 Da, respectively. Accordingly, the deuteration ratios of d-GFP$_{C0}$ and d-GFP$_{C3}$ were calculated to be 95.6 ± 1.5% and 92.5 ± 1.7%, respectively.

3.2. pD dependence of the absorption spectra

The absorption spectrum of d-GFP$_{C3}$ at pD 7.4 has two absorption peaks at 389 and 489 nm in the visible region (Supplementary Fig. S2). The wavelengths of these absorption peaks are similar to those for h-GFP$_{C0}$ and h-GFP$_{C3}$ reported previously within 5 nm (Heim *et al.*, 1995; Kneen *et al.*, 1998; Elsliger *et al.*, 1999; Takaba *et al.*, 2019). The peak heights have a pD dependence (Supplementary Fig. S2). The population of the B form increases as the pD increases (Fig. 2*b*). According to the pD dependence, the pK_a of d-GFP$_{C3}$ in D_2O was determined to be 6.64 ± 0.05, while the pK_a of h-GFP$_{C3}$ in H_2O was 5.93 ± 0.03. A ΔpK_a of ∼0.7 units on deuteration is comparable to the values reported for various small molecules (Mora-Diez *et al.*, 2015). The pK_a of h-GFP$_{C3}$ in D_2O was 6.69 ± 0.05, indicating that the increase in the pK_a is owing to replacement of the solvent. The pK_a value of h-GFP$_{C3}$ in H_2O agrees with those in previous reports for h-GFP$_{C0}$ or the S65T variant (Kneen *et al.*, 1998; Elsliger *et al.*, 1999), despite the introduction of the cycle3 mutation.



(*a*)



(*b*)

Figure 2
Properties of d-GFP$_{C3}$. (*a*) The yield of d-GFP$_{C3}$ (n = 11) is compared with those of h-GFP$_{C0}$ (n = 3), d-GFP$_{C0}$ (n = 3) and h-GFP$_{C3}$ (n = 3). Error bars indicate the standard deviation of replicates. (*b*) The pD dependence of the population ratio of the B form for d-GFP$_{C3}$ in D_2O. The estimated ratios from the absorption spectra are plotted against the pD values as green filled circles. Error bars indicate the standard deviation of triple measurements. The pH/pD dependencies for h-GFP$_{C3}$ in D_2O and h-GFP$_{C3}$ in H_2O are also shown in blue and grey, respectively.

Table 2
Refinement statistics.

	d-GFP$_{C3}$ (pD 8.5)	d-GFP$_{C3}$ (pD 7.0)	h-GFP$_{C3}$ (pD 8.5)
Resolution range (Å)	46.20–0.90	26.59–0.80	25.47–0.85
No. of reflections, working set	150858	217633	179975
No. of reflections, test set	7834	11370	9384
R_{work}† (%)	10.7	10.6	10.3
R_{free}‡ (%)	12.4	12.0	11.8
Cruickshank DPI§ (Å)	0.027	0.020	0.023
No. of non-H atoms¶			
Protein	1812.2	1806.5	1803.1
Water	461.0	511.6	500.5
No. of H atoms¶			
Protein	1530.7	1531.0	1497.7
Modelled ratio (%)	85.3	85.3	83.5
Water	43.0	38.7	49.2
R.m.s.d. from ideal			
Bond lengths (Å)	0.014	0.016	0.016
Angles (°)	2.4	2.4	2.3
Average B factors (Å^2)			
Protein	6.8	7.2	7.0
Water	21.4	26.8	26.5
Mean anisotropy††			
Protein	0.51	0.49	0.48
Water	0.40	0.34	0.35
Ramachandran plot			
Most favoured (%)	98.65	97.76	97.76
Allowed (%)	1.35	2.24	2.24
R.m.s.d. from h-GFP$_{C3}$‡‡ (Å)	0.10	0.09	0.09
PDB code	6kkz	6kl0	6kl1

† $R_{work} = \sum_{hkl} ||F_{obs}| - |F_{calc}|| / \sum_{hkl} |F_{obs}|$. ‡ R_{free} was calculated by using 5% of the reflections that were not included in the refinement as a test set. § Diffraction precision index (DPI) values were calculated using the formula $\sigma(r, B_{avg}) = 3^{1/2}(N_t/p)^{1/2}C^{-1/3}R_{free}d_{min}$ (Cruickshank, 1999). ¶ The number of atoms was calculated as the sum of occupancies. †† Anisotropy is defined as the ratio of the smallest to the largest eigenvalue of the anisotropic displacement parameter matrix (Merritt, 1999). ‡‡ The root-mean-square deviation (r.m.s.d.) was calculated with h-GFP$_{C3}$ (PDB entry 6jgi; Takaba *et al.*, 2019).

3.3. X-ray analysis at sub-ångström resolution

For the X-ray analysis of d-GFP$_{C3}$, we selected a pD value of 8.5, since we had previously elucidated the high-resolution (0.85 Å) structure of h-GFP$_{C3}$ at pH 8.5 (Takaba *et al.*, 2019). The X-ray absorption dose for each crystal position was suppressed to 2 × 10^4 Gy, which was three orders of magnitude smaller than the conventional dose limit of 2–3 × 10^7 Gy (Henderson, 1990; Owen *et al.*, 2006). The doses in this study do not cause X-ray damage to GFP and its homologous proteins (Adam *et al.*, 2009; Royant & Noirclerc-Savoye, 2011; Clavel *et al.*, 2016; Takaba *et al.*, 2019). Consequently, small differences between h-GFP$_{C3}$ and d-GFP$_{C3}$ owing to deuteration effects can be detected. The achieved resolution

for d-GFP$_{C3}$ (pD 8.5) was 0.90 Å, which was comparable to that for h-GFP$_{C3}$ (pH 8.5). The R_{work} and R_{free} factors of the final model were 10.7% and 12.4%, respectively (Table 2). Individual atoms were separately observed in the electron-density map (Supplementary Fig. S3*a*). The structures of d-GFP$_{C3}$ (pD 7.0) and h-GFP$_{C3}$ (pD 8.5) were also determined in the same way (Supplementary Figs. S3*b* and S3*c*). Refinement statistics for these analyses are listed in Table 2. For all structures, the root-mean-square deviation (r.m.s.d.) values for C^{α} atoms from h-GFP$_{C0}$ (pH 8.5) were ∼0.1 Å. This indicates that the backbone structures are almost identical despite full deuteration.

3.4. Visualization of deuterium atoms using X-rays

Almost all of the deuterium atoms located in the internal region of the structure of d-GFP$_{C3}$ (pD 8.5) were observed (Fig. 3*a*). Unobservable deuterium atoms were only located on the external surface of the molecule. In total, 85% of all possible deuterium atoms were visualized and included in the structure. Visualized hydrogen atoms of the residues on the surface were observed less often than those of the residues in the internal part of the GFP structures (Supplementary Fig. S4). Tyr66 in the chromophore proved to be deprotonated in d-GFP$_{C3}$ (pD 8.5) based on the absence of a deuterium atom at O^{η} in the electron-density map (Fig. 3*b*). Both of the two deuterium atoms of Wat3 were also visualized. One of the deuterium atoms forms a hydrogen bond to the deprotonated Tyr66, while the other forms a hydrogen bond to the main-chain carbonyl of Asn146. The deuterium atom of the hydroxyl group of Thr65 forms a hydrogen bond to N^2 of the imidazolinone ring of the chromophore (Fig. 3*b*). The carboxy group of Glu222 is observed to be neutral, while the occupancy of the major conformation is 0.78. The carboxy group

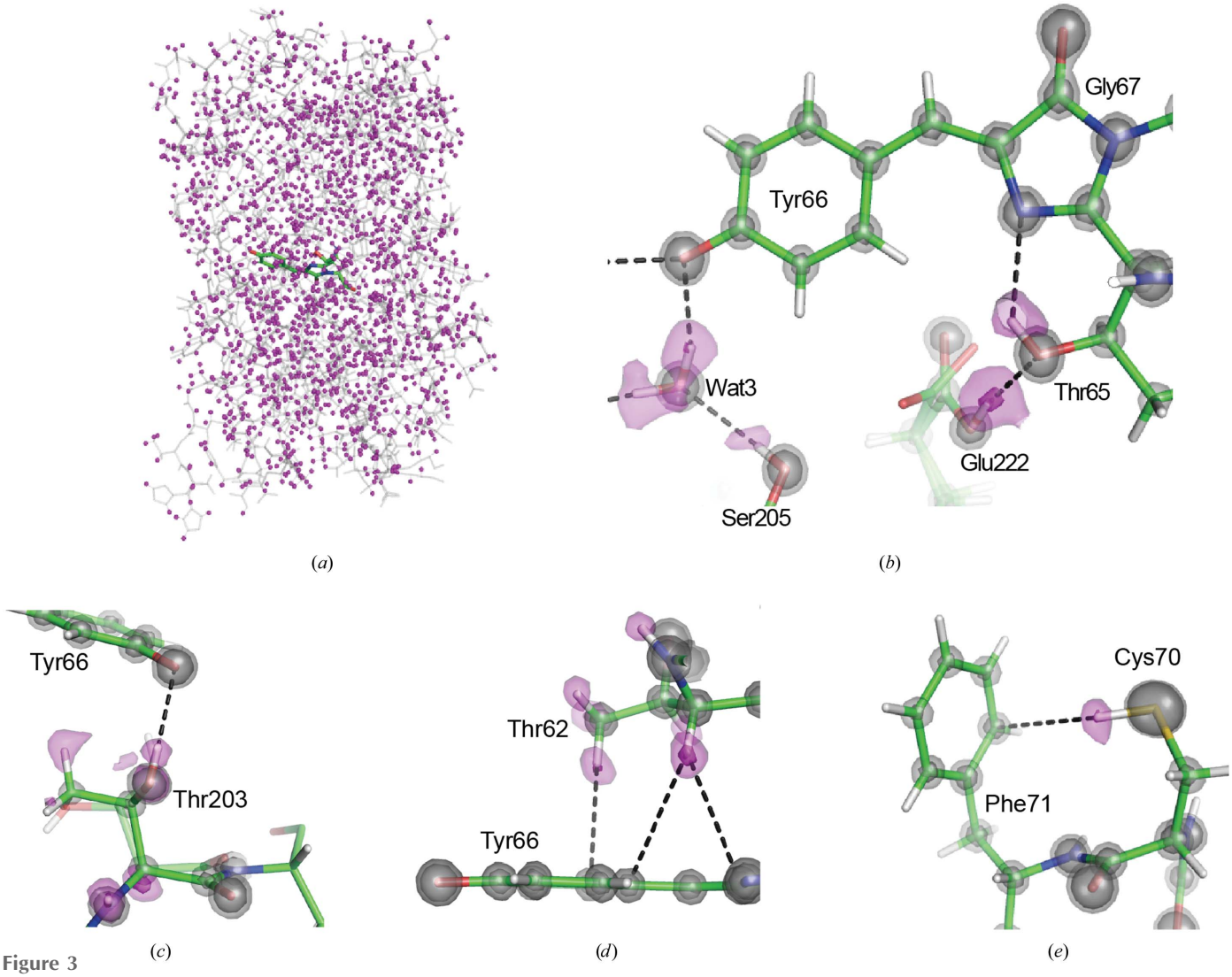


Figure 3
Visualization of deuterium atoms. (*a*) The deuterium atoms included in the model of d-GFP$_{C3}$ (pD 8.5) are shown as pink spheres. (*b*) Electron-density maps around the chromophore. The $2F_{obs} - F_{calc}$ map (5σ and 7σ levels) is shown in grey, while the $F_{obs} - F_{calc}$ hydrogen OMIT map (2σ and 4σ levels) is shown in pink. Hydrogen bonds are indicated as black dashed lines. (*c*) Electron-density maps around Thr203. The contour levels of the $F_{obs} - F_{calc}$ hydrogen OMIT map are set at 3σ and 5σ levels. (*d*) Electron-density maps around Thr62. The contour levels of the $F_{obs} - F_{calc}$ hydrogen OMIT map are set at 4σ and 6σ levels. (*e*) Electron-density map around Cys70. The contour levels of the $F_{obs} - F_{calc}$ hydrogen OMIT map are set at 4σ and 6σ levels.

interacts with O^{γ} of Thr65. Thr203 was also observed to adopt double conformations, in which the occupancy of the major conformation is 0.88. The major conformation makes a hydrogen bond to Tyr66 (Fig. 3*c*). CD–π and SD–π interactions were also clearly observed in d-GFP$_{C3}$. The methyl and methylene groups of Thr62 contact the chromophore through CD–π-type interactions (Fig. 3*d*). In the case of h-GFP$_{C3}$, these were shown to be attractive interactions according to a noncovalent interaction analysis of the charge-density distribution (Takaba *et al.*, 2019). These interactions may be conserved in d-GFP$_{C3}$ according to the absence of structural differences. Cys70 was observed to be protonated from the hydrogen OMIT map (Fig. 3*e*). The deuterium atom at S^{γ} of Cys70 interacts with Phe71. This type of interaction has been reported for some proteins as well as small molecules (Forbes *et al.*, 2017).

3.5. Overall structure of d-GFP$_{C3}$

The structures of d-GFP$_{C3}$ (pD 8.5) and h-GFP$_{C3}$ (pH 8.5) can be superimposed with a very small r.m.s.d. of 0.1 Å for C^{α} atoms (Fig. 4*a*). For some residues, the ω torsion angles (C^{α}—C—N′—$C^{\alpha\prime}$) of the peptide bonds are observed to deviate from the planar structure with $\omega = 180°$. The smallest ω value of 158° was observed for the Arg96–Thr97 peptide bond (Fig. 4*b*). It is considered that the hydrogen bond between the chromophore and Arg96 causes this distortion of the peptide bond between Arg96 and Thr97. The largest value of 200° was at Gly40–Lys41 (Fig. 4*c*). The peptide bond between Gly40 and Lys41 is located at the terminus of a strand. Similarly, almost distorted residues with $|\omega - 180°|$ values larger than 10° interact with the chromophore or are at the termini of strands. The differences in the ω values between h-GFP$_{C3}$ and d-GFP$_{C3}$ are lower than 5° even for these peptide bonds (Fig. 4*d*). The residues interacting with the chromophore have large distortions of the peptide bonds, as observed in h-GFP$_{C3}$ (Takaba *et al.*, 2019). The *B*-factor values of central residues and the chromophore were observed to be lower than those of external residues in the structure of d-GFP$_{C3}$ (pD 8.5) (Supplementary Fig. S5). Similar distributions of *B* factors were also observed in the structures of d-GFP$_{C3}$ (pD 7.0) and h-GFP$_{C3}$ (pD 8.5).

3.6. Structures of the chromophore of d-GFP$_{C3}$

Accurate geometric information for the chromophore was obtained because it is in the central portion of the molecule (Supplementary Table S1). Since the bond length of the bridge bond $C^{\beta 2}$—$C^{\gamma 2}$ is highly correlated with the atomic charge on O^{η} (Weber *et al.*, 1999), a plot of the C^{ζ}—O^{η} and $C^{\beta 2}$—$C^{\gamma 2}$ bond lengths was made in order to compare them with various crystallographic results at high resolution (Fig. 5*a*). The C^{ζ}—O^{η} bond length is 1.311 ± 0.009 Å for d-GFP$_{C3}$ (pD 8.5), which is intermediate between the values for single (d_{C-O} = ∼1.38 Å) and double ($d_{C=O}$ = ∼1.24 Å) bonds. This value is nearly identical (within the margin of error) to that for h-GFP$_{C3}$ in H_2O (pH 8.5). In addition, d-GFP$_{C3}$ (pD 7.0) and h-GFP$_{C3}$ (pD 8.5) are also found within the same area of the plot. This indicates that the electronic structures of these proteins are essentially identical despite some differences in the deuteration condition. This is supported by the observations that deuteration provides no changes in the absorption

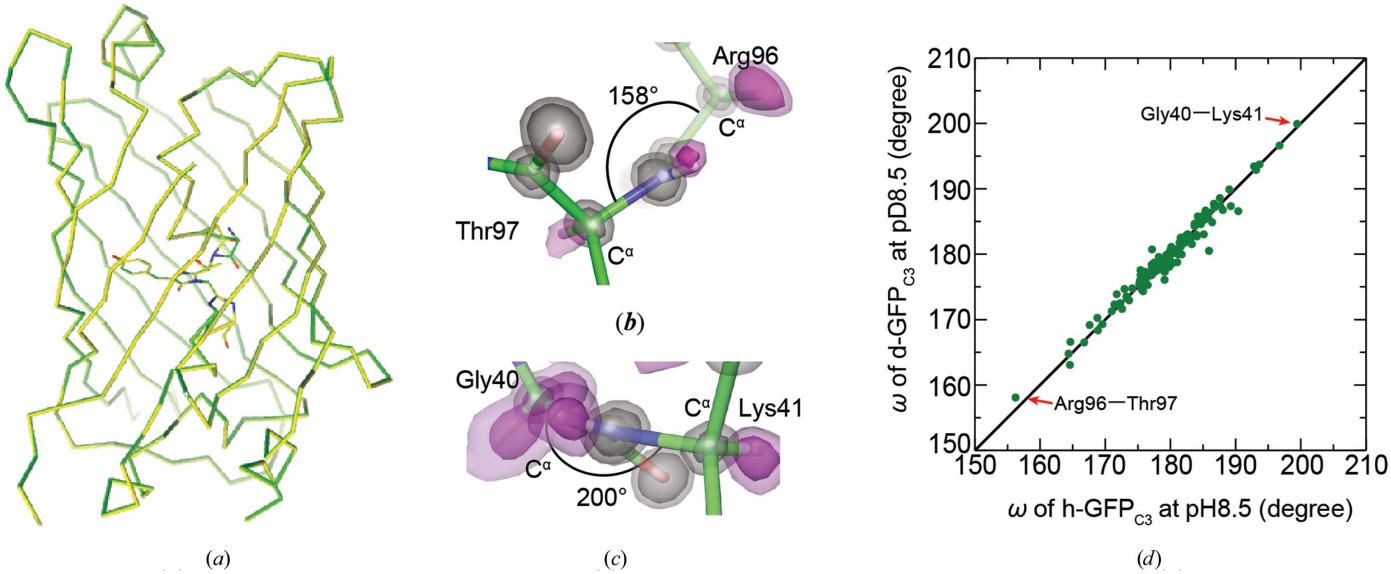


Figure 4
Analyses of the main-chain structure. (*a*) The C^{α} model of d-GFP$_{C3}$ (pD 8.5) and that of h-GFP$_{C3}$ (pH 8.5) are superimposed and shown in green and yellow, respectively. The chromophores for these proteins are also indicated as stick models. (*b*) The distortion of the peptide bond between Arg96 and Thr97 with an ω torsion angle of 158°. The $2F_{obs} - F_{calc}$ map (5σ and 7σ levels) is shown in grey, while the $F_{obs} - F_{calc}$ hydrogen OMIT map (4σ and 5σ levels) is shown in pink. (*c*) The distortion of the peptide bond between Gly40 and Lys41 with an ω torsion angle of 200°. The $2F_{obs} - F_{calc}$ map (5σ and 7σ levels) is shown in grey, while the $F_{obs} - F_{calc}$ hydrogen OMIT map (2σ and 3σ levels) is shown in pink. (*d*) The relationship between the ω torsion angles of peptide bonds in d-GFP$_{C3}$ (pD 8.5) and h-GFP$_{C3}$ (pH 8.5). Only residues in which the occupancy of the major conformation is greater than 0.7 are plotted.

Table 3
Hydrogen-bond distances (d_{D-A}, Å) around the chromophore.

Values in parentheses are estimated standard deviations derived from full-matrix least-squares refinement.

Donor	Acceptor	d-GFP$_{C3}$ (pD 8.5)	d-GFP$_{C3}$ (pD 7.0)	h-GFP$_{C3}$ (pD 8.5)	h-GFP$_{C3}$ (pH 8.5) (Takaba *et al.*, 2019)
Arg168 N	His148 N$^{\varepsilon 2}$	3.086 (10)	3.164 (9)	3.082 (9)	3.157 (9)
His148 N$^{\delta 1}$	Tyr66 O$^{\eta}$	2.872 (9)	2.866 (7)	2.871 (8)	2.869 (8)
Wat3	Tyr66 O$^{\eta}$	2.719 (9)	2.735 (6)	2.728 (8)	2.737 (7)
Wat3	Asn146 O	2.873 (10)	2.892 (7)	2.886 (8)	2.903 (7)
Ser205 O$^{\gamma}$	Wat3	2.757 (9)	2.760 (7)	2.767 (8)	2.762 (7)
Glu222 O$^{\varepsilon 2}$	Thr65 O$^{\gamma 1}$	2.681 (8)†	2.686 (6)†	2.685 (7)†	2.670 (6)†
Thr65 O$^{\gamma 1}$	Tyr66 N^{2}	2.745 (9)	2.737 (7)	2.740 (8)	2.740 (7)
Thr203 O$^{\gamma}$	Tyr66 O$^{\eta}$	2.668 (8)†	2.687 (6)†	2.665 (6)†	2.672 (6)†

† These values are for the major alternative conformation.

and fluorescent wavelengths. On the other hand, other B-form structures are widely scattered in the lower area of the plot, indicating that the C$^{\zeta}$—O$^{\eta}$ bond length has a large variety. This implies that the bond lengths in the chromophore in previous studies are strongly influenced by the geometric restraints as single or double bonds or by significant X-ray damage.

3.7. Comparison of hydrogen-bond distances

The hydrogen-bond distances around the chromophore as well as those between the carbonyl and amide in the main chain were compared between h-GFP$_{C3}$ (pH 8.5) and d-GFP$_{C3}$ (pD 8.5) (Fig. 5*b*). It should be noted that the hydrogen-bond distances in d-GFP$_{C3}$ which are longer than 3 Å are generally shorter than those in h-GFP$_{C3}$. On the other hand, the hydrogen-bond distances in d-GFP$_{C3}$ which are shorter than 3 Å are similar to those in h-GFP$_{C3}$. For hydrogen bonds around the chromophore, two hydrogen-bond distances between the main-chain carbonyl of Asn146 and Wat3 and between N$^{\varepsilon 2}$ of His148 and the amide N atom of Arg168 have significant differences larger than the margin of error (Table 3). These observations may suggest that the deuterium atoms form stronger bonds. However, it has been shown that the enhanced solvent–solvent interaction of D_2O plays a critical role in the higher stability of many proteins in D_2O (Parker & Clarke, 1997; Cioni & Stambini, 2002). In fact, the hydrogen bonds in D_2O have a deeper potential curve for hydrogen-bond dissociation (Sheu *et al.*, 2008). That is, the hydrogen bonds are strengthened not only by direct effects at the bonds themselves but also by numerous indirect factors, including the solvent.

3.8. Protonation state of His148

In d-GFP$_{C3}$ (pD 8.5), the deuterium atom at N$^{\varepsilon 2}$ of His148 is not observed in the electron-density map (Fig. 6*a*). This

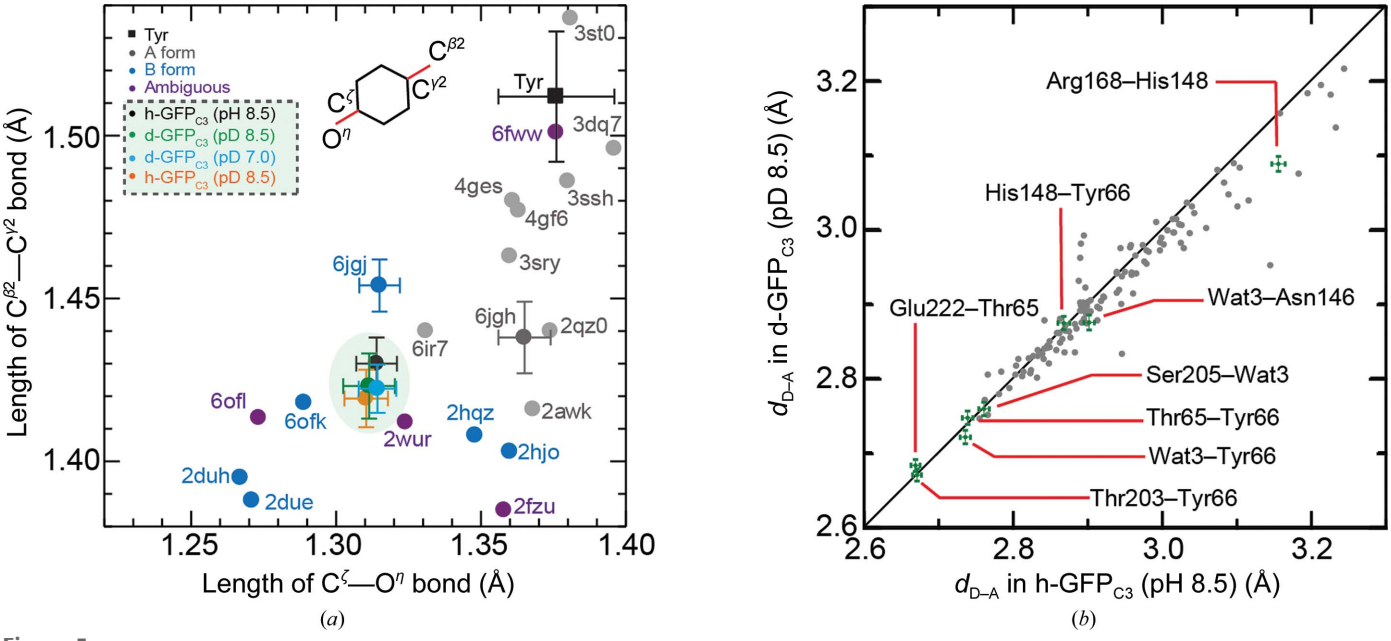


Figure 5
Comparisons of geometric parameters around the chromophore. (*a*) The relationship between the C$^{\zeta}$—O$^{\eta}$ and C$^{\beta 2}$—C$^{\gamma 2}$ bond lengths in the chromophore are plotted for various GFP structures with resolutions of better than 1.3 Å (Supplementary Table S2). Grey filled circles indicate variants considered to adopt the A-form structure, while blue filled circles indicate variants considered to adopt the B-form structure. Those with ambivalent interpretations are shown in purple. The C$^{\zeta}$—O$^{\eta}$ and C$^{\beta 2}$—C$^{\gamma 2}$ bond lengths of the structures in this work and our previous work (Takaba *et al.*, 2019) are also plotted. The error bars are the standard deviations calculated by full-matrix refinement with *SHELXL*. The bond lengths for tyrosine are also indicated as a filled black square with standard deviations for the restraint. (*b*) The relationships between hydrogen bonds of d-GFP$_{C3}$ (pD 8.5) and h-GFP$_{C3}$ (pH 8.5). The plots for hydrogen bonds formed between the chromophore and the surrounding residues are coloured green, while those formed between the carbonyl and the amide in the main chain are coloured grey. The error bars are the standard deviations calculated by full-matrix refinement with *SHELXL*.

indicates that His148 is deprotonated at pD 8.5. On the other hand, our previous report indicated that His148 in h-GFP$_{C3}$ (pH 8.5) is protonated (Takaba *et al.*, 2019). In order to obtain more information about the protonation of His148, we further compared the d-GFP$_{C3}$ structure with that at pD 7.0. Electron density for the deuterium atom was clearly observed at N$^{\varepsilon 2}$ of His148 in the electron-density map (Fig. 6*b*). The deuterium atom at the main-chain amide N atom of Arg168 was observed simultaneously, and its proximity to the deuterium atom at N$^{\varepsilon 2}$ of His148 may be unfavourable for interaction. In the case of h-GFP$_{C3}$ (pD 8.5), His148 is deprotonated as in d-GFP$_{C3}$ (pD 8.5) (Fig. 6*c*). Taken together with these OMIT maps, the pK_a of His148 is decreased in D_2O. The protonation at His148 may be influenced not by nondissociable hydrogens but rather by dissociable hydrogens in GFP$_{C3}$. The distance between N$^{\varepsilon 2}$ of His148 and the amide N atom of Arg168 for d-GFP$_{C3}$ (pD 7.0) is 3.164 ± 0.009 Å (Table 3). This distance is almost identical to that for h-GFP$_{C3}$ (pH 8.5) at 3.157 ± 0.009 Å, while those for d-GFP$_{C3}$ (pD 8.5) and h-GFP$_{C3}$ (pD 8.5) are 3.086 ± 0.010 and 3.082 ± 0.009 Å, respectively. The σ level for the distance between d-GFP$_{C3}$ (pD 8.5) and h-GFP$_{C3}$ (pH 8.5) is estimated to be 5.3. This indicates that the difference between the two is meaningful but small.

4. Discussion

We investigated the hydrogen isotope effects on GFP by comparing them between h-GFP$_{C3}$ in H_2O and d-GFP$_{C3}$ in D_2O using crystallographic and spectroscopic procedures. Even at sub-ångström resolutions, the X-ray structures of the two proteins showed no apparent differences in chromophore geometry or overall structure. However, the protonation state of His148 near the chromophore in d-GFP$_{C3}$ (pD 8.5) was different from that in h-GFP$_{C3}$ (pH 8.5). Because h-GFP$_{C3}$ (pD 8.5) shows the same protonation state at His148 as d-GFP$_{C3}$ (pD 8.5), the difference is owing to dissociable hydrogens in GFP$_{C3}$. However, His148 of d-GFP$_{C3}$ is protonated at pD 7.0, while the chromophore is still deprotonated.

To date, the X-ray structures of several proteins have been investigated for both protiated and perdeuterated samples. In almost all of these, no meaningful differences were observed in the structures (Gamble *et al.*, 1994; Meilleur *et al.*, 2005; Artero *et al.*, 2005; Wang *et al.*, 2006). Especially for the cases of human aldose reductase and high-potential iron–sulfur protein from *Thermochromatium tepidum*, no significant differences were detected by the X-ray analyses, which were performed accurately by including almost all of the H atoms (Blakeley *et al.*, 2008; Hanazono *et al.*, 2019). On the other hand, haloalkane dehydrogenase from *Xanthobacter autotrophicus* exhibited large conformational differences between protiated and deuterated samples despite a medium resolution of ∼1.5 Å (Liu *et al.*, 2007). However, the pH/pD values of the crystallization conditions differed significantly between the protiated and deuterated samples. Therefore, our result is a valuable example in which structural perturbations in hydrogen bonding near the active site of the protein are detected on deuteration.

Spectroscopically silent as well as spectroscopically active pK_a changes on deuteration were detected around the chromophore in this study. The pD values were corrected from those measured using the pH meter by adding by 0.4 units (Glasoe & Long, 1960), while the difference between the true and measured values was indicated to be negligible or small in a recent paper (Rubinson, 2017). Even if we use the measured values without any corrections in this study, the conclusion that the pK_a of His148 decreases after changing the solvent from H_2O to D_2O remains unchanged. The pK_a values of residues in proteins differ from those of free amino acids in solution because of various interactions in the protein (Ma *et al.*, 1999). In the case of h-GFP$_{C3}$, the pK_a of Tyr66 is regulated by hydrogen bonds to surrounding residues such as His148, Thr203 and Wat3 (Brejc *et al.*, 1997). Consequently, Tyr66, a component residue of the chromophore, has a significantly lower pK_a of ∼6 than that for free tyrosine (10.5). This shift cannot be explained only by the π-resonance of the chromophore as the pK_a of a chromophore analogue, 4-hydroxybenzylidene-imidazolinone, is 8.3 (Scharnagl & Raupp-Kossmann, 2004). Our structural analyses imply that the positive charge of His148 plays an important role in the stabilization of the negative charge of the deprotonated Tyr66. Indeed, it has been reported that the pK_a of Tyr66 increases when His148 is replaced by neutral or acidic residues in the S65T variant (Shu *et al.*, 2007). Therefore, it is suggested that the positive charge of His148 is important for accumulating

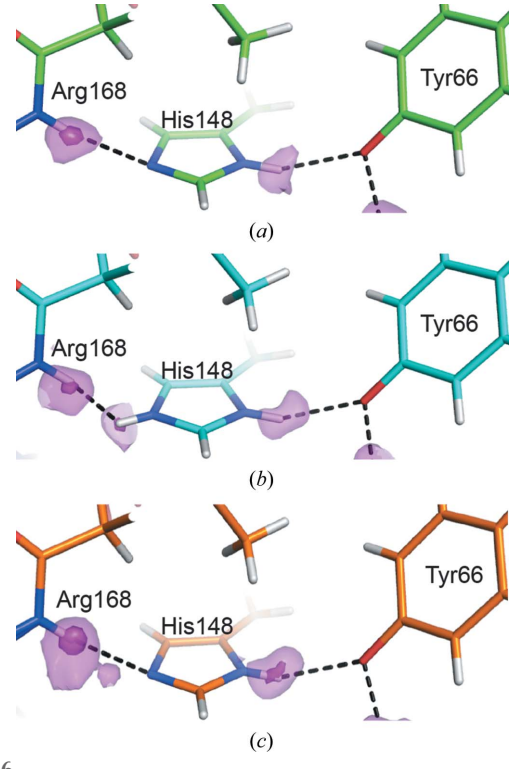


Figure 6
Protonation states at His148. (*a*) The $F_{obs} - F_{calc}$ hydrogen OMIT map (2σ and 4σ levels) around His148 of d-GFP$_{C3}$ (pD 8.5) is shown in pink. (*b*) The map for d-GFP$_{C3}$ (pD 7.0). (*c*) The map for h-GFP$_{C3}$ (pD 8.5).

the light-emitting deprotonated form of the chromophore at the physiological pH of ∼7.

In usual cases, the pK_a in D_2O is higher than that in H_2O. However, the pK_a of His148 in D_2O seems to be lower than that in H_2O. This result may be explained by proton sharing, which can exhibit significant isotope effects. For the case of h-GFP$_{C3}$ (pH 8.5), a proton may be shared between N$^{\varepsilon 2}$ of His148 and the amide N atom of Arg168, because the latter is deprotonated and has a negative charge owing to the basic conditions. In the case of d-GFP$_{C3}$ (pD 8.5) this sharing is more difficult owing to the higher pK_a of the amide N atom of Arg168 in the deuterated sample. Therefore, the deuterium atom is localized at Arg168 in d-GFP at pD 8.5. Similar proton-shared structures have been reported for cholesterol oxidase at pH 9.0 (Lyubimov *et al.*, 2006; Golden *et al.*, 2017). Certainly, we can suggest some alternative interpretations: for example, H atoms may simultaneously coexist at both sites. However, there is only small possibility of this interpretation for the following reason. If the NH of the amide of Arg168 is not deprotonated and His148 is protonated, the side chain of His148 would rotate owing to repulsive force. Otherwise, the distance between the N$^{\varepsilon 2}$ atom of His148 and the amide N atom of Arg168 would become larger than normal hydrogen-bond distances. In the future, we will investigate this problem using neutron crystallography, since the complicated manner of protonation is expected to be determined in greater detail using this method. Neutron analyses of d-GFP$_{C3}$ should be carried out at pD 7.0 in addition to pD 8.5 in order to understand the strange protonation state of His148.

In conclusion, this study demonstrates that protonation states can differ between deuterated and protiated samples. Therefore, we should carefully investigate the differences in the protonation states in the vicinity of the active site prior to the use of deuterated proteins in neutron crystallographic studies. This may be critical in order to reach any reliable conclusions from neutron crystallographic analyses using deuterated proteins.

Acknowledgements

We thank Dr S. Niwa for her critical reading of the manuscript. We also thank the beamline staff at BL41XU and BL44XU of SPring-8 for their help with the diffraction experiments (Proposal Nos. 2015B2033, 2015B1037, 2016A6624, 2016A2536, 2016B6624, 2017A6728, 2017A2542, 2017B6728, 2018A6827, 2018B2705 and 2019A2560 to KaT). Author contributions were as follows. KaT supervised the project, YT and KaT designed the experiments, YT, KiT and YH established the expression, purification and crystallization protocols, YT prepared the crystals for X-ray data collection, YT, YH and KaT performed the X-ray data collections, YT and HAD performed spectroscopic analyses, YT, KiT, YH, HAD and KaT performed the crystallographic analysis, YT, KiT, YH, HAD, KM and KaT discussed the results, YT wrote the initial draft and KaT revised it. All authors gave comments on the manuscript and consented to the submission of the final version.

Funding information

This work was supported by a JSPS KAKENHI grant (No. JP17H03643 to KaT), by funds from the Ishizue Grant Program of Kyoto University (to KaT) and by a grant from the Photon and Quantum Basic Research Coordinated Development Program of the Ministry of Education, Culture, Sports, Science and Technology of Japan (to KM).

References

Adam, V., Carpentier, P., Violot, S., Lelimousin, M., Darnault, C., Nienhaus, G. U. & Bourgeois, D. (2009). *J. Am. Chem. Soc.* **131**, 18063–18065.
Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C. & Zwart, P. H. (2010). *Acta Cryst.* D**66**, 213–221.
Afonine, P. V., Mustyakimov, M., Grosse-Kunstleve, R. W., Moriarty, N. W., Langan, P. & Adams, P. D. (2010). *Acta Cryst.* D**66**, 1153–1163.
Artero, J.-B., Härtlein, M., McSweeney, S. & Timmins, P. (2005). *Acta Cryst.* D**61**, 1541–1549.
Ashkar, R., Bilheux, H. Z., Bordallo, H., Briber, R., Callaway, D. J. E., Cheng, X., Chu, X.-Q., Curtis, J. E., Dadmun, M., Fenimore, P., Fushman, D., Gabel, F., Gupta, K., Herberle, F., Heinrich, F., Hong, L., Katsaras, J., Kelman, Z., Kharlampieva, E., Kneller, G. R., Kovalevsky, A., Krueger, S., Langan, P., Lieberman, R., Liu, Y., Losche, M., Lyman, E., Mao, Y., Marino, J., Mattos, C., Meilleur, F., Moody, P., Nickels, J. D., O'Dell, W. B., O'Neill, H., Perez-Salas, U., Peters, J., Petridis, L., Sokolov, A. P., Stanley, C., Wagner, N., Weinrich, M., Weiss, K., Wymore, T., Zhang, Y. & Smith, J. C. (2018). *Acta Cryst.* D**74**, 1129–1168.
Blakeley, M. P., Hasnain, S. S. & Antonyuk, S. V. (2015). *IUCrJ*, **2**, 464–474.
Blakeley, M. P., Ruiz, F., Cachau, R., Hazemann, I., Meilleur, F., Mitschler, A., Ginell, S., Afonine, P., Ventura, O. N., Cousido-Siah, A., Haertlein, M., Joachimiak, A., Myles, D. & Podjarny, A. (2008). *Proc. Natl Acad. Sci. USA*, **105**, 1844–1848.
Brejc, K., Sixma, T. K., Kitts, P. A., Kain, S. R., Tsien, R. Y., Ormö, M. & Remington, S. J. (1997). *Proc. Natl Acad. Sci. USA*, **94**, 2306–2311.
Brockwell, D., Yu, L., Cooper, S., McCleland, S., Cooper, A., Attwood, D., Gaskell, S. J. & Barber, J. (2001). *Protein Sci.* **10**, 572–580.
Chattoraj, M., King, B. A., Bublitz, G. U. & Boxert, S. G. (1996). *Proc. Natl Acad. Sci. USA*, **93**, 8362–8367.
Chen, C.-H., Liu, I.-W., MacColl, R. & Berns, D. (1984). *Biopolymers*, **22**, 1223–1233.
Chen, J. C.-H. & Unkefer, C. J. (2017). *IUCrJ*, **4**, 72–86.
Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* D**66**, 12–21.
Cioni, P. & Stambini, G. B. (2002). *Biophys. J.* **82**, 3246–3253.
Clavel, D., Gotthard, G., von Stetten, D., De Sanctis, D., Pasquier, H., Lambert, G. G., Shaner, N. C. & Royant, A. (2016). *Acta Cryst.* D**72**, 1298–1307.
Cleland, W. W. (2005). *Arch. Biochem. Biophys.* **433**, 2–12.
Cruickshank, D. W. J. (1999). *Acta Cryst.* D**55**, 583–601.
DeLano, W. L. (2002). *PyMOL*. http://www.pymol.org.
Elsliger, M. A., Wachter, R. M., Hanson, G. T., Kallio, K. & Remington, S. J. (1999). *Biochemistry*, **38**, 5296–5301.
Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.
Flot, D., Mairs, T., Giraud, T., Guijarro, M., Lesourd, M., Rey, V., van Brussel, D., Morawe, C., Borel, C., Hignette, O., Chavanne, J.,

Nurizzo, D., McSweeney, S. & Mitchell, E. (2010). *J. Synchrotron Rad.* **17**, 107–118.
Forbes, C. R., Sinha, S. K., Ganguly, H. K., Bai, S., Yap, G. P. A., Patel, S. & Zondlo, N. J. (2017). *J. Am. Chem. Soc.* **139**, 1842–1855.
Fukuda, H., Arai, M. & Kuwajima, K. (2000). *Biochemistry*, **39**, 12025–12032.
Gamble, T. R., Clauser, K. R. & Kossiakoff, A. A. (1994). *Biophys. Chem.* **53**, 15–25.
Glasoe, P. K. & Long, F. A. (1960). *J. Phys. Chem.* **64**, 188–190.
Golden, E., Yu, L.-J., Meilleur, F., Blakeley, M. P., Duff, A. P., Karton, A. & Vrielink, A. (2017). *Sci. Rep.* **7**, 40517.
Gu, H. & Zhang, S. (2013). *Molecules*, **18**, 9278–9292.
Hanazono, Y., Takeda, K. & Miki, K. (2019). *Proteins*, https://doi.org/10.1002/prot.25793.
Hattori, A., Crespi, H. L. & Katz, J. J. (1965). *Biochemistry*, **4**, 1213–1225.
Hazemann, I., Dauvergne, M. T., Blakeley, M. P., Meilleur, F., Haertlein, M., Van Dorsselaer, A., Mitschler, A., Myles, D. A. A. & Podjarny, A. (2005). *Acta Cryst.* D**61**, 1413–1417.
Heim, R., Cubitt, A. B. & Tsien, R. Y. (1995). *Nature (London)*, **373**, 663–664.
Heim, R. & Tsien, R. Y. (1996). *Curr. Biol.* **6**, 178–182.
Henderson, R. (1990). *Proc. R. Soc. London Ser. B*, **241**, 6–8.
Hohlefelder, L. S., Stögbauer, T., Opitz, M., Bayerl, T. M. & Rädler, J. O. (2013). *Biomed. Res. Int.*, **2013**, 592745.
Ishikita, H. & Saito, K. (2013). *J. R. Soc. Interface*, **11**, 20130518.
Karplus, P. A. & Diederichs, K. (2012). *Science*, **336**, 1030–1033.
Kneen, M., Farinas, J., Li, Y. & Verkman, A. S. (1998). *Biophys. J.* **74**, 1591–1599.
Kummer, A. D., Kompa, C., Lossau, H., Pöllinger-Dammer, F., Michel-Beyerle, M. E., Silva, C. M., Bylina, E. J., Coleman, W. J., Yang, M. M. & Youvan, D. C. (1998). *Chem. Phys.* **237**, 183–193.
Lins, L., Thomas, A. & Brasseur, R. (2003). *Protein Sci.* **12**, 1406–1417.
Liu, X., Hanson, B. L., Langan, P. & Viola, R. E. (2007). *Acta Cryst.* D**63**, 1000–1008.
Lyubimov, A. Y., Lario, P. I., Moustafa, I. & Vrielink, A. (2006). *Nature Chem. Biol.* **2**, 259–264.
Ma, K., Clancy, E. L., Zhang, Y., Ray, D. G., Wollenberg, R. & Zagorski, M. G. (1999). *J. Am. Chem. Soc.* **121**, 8698–8706.
McAnaney, T. B., Park, E. S., Hanson, G. T., Remington, S. J. & Boxer, S. G. (2002). *Biochemistry*, **41**, 15489–15494.
Meilleur, F., Dauvergne, M.-T., Schlichting, I. & Myles, D. A. A. (2005). *Acta Cryst.* D**61**, 539–544.
Merritt, E. A. (1999). *Acta Cryst.* D**55**, 1109–1117.
Mora-Diez, N., Egorova, Y., Plommer, H. & Tremaine, P. R. (2015). *RSC Adv.* **5**, 9097–9109.
Morris, M. C. (2013). *Biochim. Biophys. Acta*, **1834**, 1387–1395.
Nickels, J. D., O'Neill, H., Hong, L., Tyagi, M., Ehlers, G., Weiss, K. L., Zhang, Q., Yi, Z., Mamontov, E., Smith, J. C. & Sokolov, A. P. (2012). *Biophys. J.* **103**, 1566–1575.
Nickels, J. D., Perticaroli, S., O'Neill, H., Zhang, Q., Ehlers, G. & Sokolov, A. P. (2013). *Biophys. J.* **105**, 2182–2187.
Oksanen, E., Chen, J. C.-H. & Fisher, S. Z. (2017). *Molecules*, **22**, 596–631.
Ormö, M., Cubitt, A. B., Kallio, K., Gross, L. A., Tsien, R. Y. & Remington, S. J. (1996). *Science*, **273**, 1392–1395.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Owen, R. L., Rudiño-Piñera, E. & Garman, E. F. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 4912–4917.
Paithankar, K. S., Owen, R. L. & Garman, E. F. (2009). *J. Synchrotron Rad.* **16**, 152–162.
Parker, M. J. & Clarke, A. R. (1997). *Biochemistry*, **36**, 5786–5794.
Piszczek, G., Lee, J. C., Tjandra, N., Lee, C.-R., Seok, Y.-J., Levine, R. L. & Peterkofsky, A. (2011). *Arch. Biochem. Biophys.* **507**, 332–342.
Royant, A. & Noirclerc-Savoye, M. (2011). *J. Struct. Biol.* **174**, 385–390.
Rubinson, K. A. (2017). *Anal. Methods*, **9**, 2744–2750.
Scharnagl, C. & Raupp-Kossmann, R. A. (2004). *J. Phys. Chem. B*, **108**, 477–489.
Schowen, K. B., Limbach, H. H., Denisov, G. S. & Schowen, R. L. (2000). *Biochim. Biophys. Acta*, **1458**, 43–62.
Schramm, V. L. (2007). *Curr. Opin. Chem. Biol.* **11**, 529–536.
Sheldrick, G. & Schneider, T. (1997). *Methods Enzymol.* **277**, 319–343.
Sheu, S.-Y., Schlag, E. W., Selzle, H. L. & Yang, D.-Y. (2008). *J. Phys. Chem. A*, **112**, 797–802.
Shimomura, O., Johnson, F. H. & Saiga, Y. (1962). *J. Cell. Comput. Physiol.* **59**, 223–239.
Shu, X., Kallio, K., Shi, X., Abbyad, P., Kanchanawong, P., Childs, W., Boxer, S. G. & Remington, S. J. (2007). *Biochemistry*, **46**, 12005–12013.
Sirerol-Piquer, M. S., Cebrián-Silla, A., Alfaro-Cervelló, C., Gomez-Pinedo, U., Soriano-Navarro, M. & Verdugo, J. M. (2012). *Micron*, **43**, 589–599.
Takaba, K., Tai, Y., Eki, H., Dao, H.-A., Hanazono, Y., Hasegawa, K., Miki, K. & Takeda, K. (2019). *IUCrJ*, **6**, 387–400.
Wang, Z., Li, C., Ellenburg, M., Soistman, E., Ruble, J., Wright, B., Ho, J. X. & Carter, D. C. (2006). *Acta Cryst.* D**62**, 800–806.
Weber, W., Helms, V., McCammon, J. A. & Langhoff, P. W. (1999). *Proc. Natl Acad. Sci. USA*, **96**, 6177–6182.
Yang, F., Moss, L. G. & Phillips, G. N. Jr (1996). *Nature Biotechnol.* **14**, 1246–1251.